Deutscher Bundestag **Drucksache 19/1484**

19. Wahlperiode 22.03.2018

Vorabfassung - wird durch die lektorierte Version ersetzt.

Kleine Anfrage

der Abgeordneten Ulla Jelpke, Dr. André Hahn, Gökay Akbulut, Anke Domscheit-Berg, Amira Mohamed Ali, Zaklin Nastic, Niema Movassat, Dr. Petra Sitte, Kersten Steinke, Friedrich Straetmanns, Dr. Kirsten Tackmann und der Fraktion DIE LINKE.

Einsatz von Spracherkennungssoftware durch das Bundesamt für Migration und Flüchtlinge

Seit 1998 gibt das Bundesamt für Migration und Flüchtlinge (BAMF) bei sprachwissenschaftlichen Gutachtern Sprach- und Textanalysen (STA) zur Verifikation der Herkunft von Asylantragstellerinnen und Asylantragstellern in Auftrag. Diese werden auf Grundlage von dreißigminütigen Interviews erstellt (Süddeutsche Zeitung vom 7. Oktober 2016, www.sueddeutsche.de/wissen/migration-die-grenzen-der-sprache-1.3193855).

Im Frühjahr 2017 wurden über Presseberichte Pläne des BAMF bekannt, zukünftig Sprachproben von Asylbewerberinnen und -bewerbern aufzunehmen und diese mittels einer speziellen Software einer automatischen Dialektanalyse zu unterziehen (Die Welt vom 17. März 2017, www.welt.de/wissenschaft/article162926845/Software-soll-Dialekt-von-Asylbewerbern-untersuchen.html). Laut BAMF soll die biometrische Sprachanalyse mittels Dialekterkennungssoftware den Entscheiderinnen und Entscheidern helfen, Angaben zum Herkunftsland zu überprüfen, wenn Asylsuchende nicht über Identitätspapiere verfügen. Im Juli 2017 informierte das BAMF darüber, dass im Rahmen eines Modellprojekts in Bamberg verschiedene technische Assistenzsysteme für die Identitätsprüfung erprobt würden, darunter auch eine Software für Akzenterkennung, die es möglich machen soll, auf Basis einer zweiminütigen Sprachprobe den gesprochenen Dialekt zu ermitteln (Aktuelle Meldungen des BAMF, 26. Juli 2017, www.bamf.de/SharedDocs/Meldungen/DE/2017/20170726-am-vorstellung-modellprojekt-bamberg.html). Ende Juli 2017 kündigte das BAMF an, die erprobten Systeme ab August 2017 innerhalb weniger Monate bundesweit einzuführen (Die Welt vom 25. Juli 2017, www.welt.de/politik/deutschland/article167043676/Neue-Software-soll-Fluechtlinge-schneller-identifizieren.html). Im Oktober 2017 teilte das BAMF mit, es seien bereits „1200 nutzbare Sprachanalysen erstellt“ worden. Seit September sei das speziell für das BAMF entwickelte System im Einsatz. Es erkenne die wichtigsten arabischen Dialekte und liefere Anhaltspunkte dafür, ob die Angaben der Flüchtlinge stimmen (Rheinische Post vom 31. Oktober 2017, www.presseportal.de/pm/30621/3775263). Laut Angaben der Bundesregierung ist ein flächendeckender Einsatz der Sprachbiometrie-Software für April 2018 geplant (vgl. Antwort auf die Kleine Anfrage der Fraktion DIE LINKE, Bundestagsdrucksache 19/190, S. 2).

Die Zuverlässigkeit solcher automatischer Analysen ist allerdings umstritten. Im Rahmen einer wissenschaftlichen Untersuchung aus dem Jahr 2016 erzielte die verwendete Software bei der Zuordnung einer Sprachaufnahme zu einem der vier

am häufigsten gesprochenen arabischen Dialekte nur eine Genauigkeit von 60 Prozent (Ali et al. 2016, www.research.ed.ac.uk/portal/files/26891896/aali_is2016.pdf). Linguistinnen und Linguisten weisen ferner darauf hin, dass sich die Sprechweise von Menschen über die Lebensspanne verändern kann und daher nicht zwingend Rückschlüsse auf ihre Herkunft erlaubt. Dies betreffe insbesondere Personen, die mit unterschiedlichen Dialekten aufgewachsen sind, weil sie beispielsweise eine jahrelange Fluchtgeschichte hinter sich haben (FAZ vom 5. März 2017, www.faz.net/aktuell/wissen/geist-soziales/linguistische-forensik-ordnet-menschen-ihrer-nationalitaet-zu-14896182.html?printPagedArticle=true#pageIndex_0). Selbst eine erfolgreiche Dialektbestimmung kann schließlich lediglich Hinweise auf die Herkunft eines Menschen geben, da Sprach- und Landesgrenzen nicht zwingend übereinstimmen (CILIP vom 23. November 2017, www.cilip.de/2017/11/23/digitalisierte-migrationskontrolle-wenn-technik-ueber-asyl-entscheidet). Das BAMF betont in diesem Zusammenhang immer wieder, dass die endgültige Entscheidung von Menschen getroffen werde. Die mithilfe der Software erlangten Erkenntnisse würden lediglich zusätzliche Anhaltspunkte liefern (vgl. ebd.). Auch die Bundesregierung hebt hervor, dass die Software lediglich „dem Erhalt eines ergänzenden Beitrags bzw. Indizes zur Verifikation des Herkunftsstaates oder der Herkunftsregion des Antragstellers“ diene (Bundestagsdrucksache 19/190, S. 5). Dies setzt voraus, dass die Entscheiderinnen und Entscheider in der Lage sind, die Ergebnisse der softwarebasierten Analysen fachkundig zu interpretieren und sie zu weiteren Anhaltspunkten in Beziehung zu setzen.

Jedoch bestehen bereits unabhängig vom Einsatz digitaler Verfahren im BAMF Mängel bei der Ausbildung der Entscheiderinnen und Entscheider. Laut einem Bericht von Zeit Online von Juni 2017 wurden im BAMF seit August 2015 mehrere Tausend neue Mitarbeiterinnen und Mitarbeiter eingestellt. Nach einer internen „Auswertung der Qualifizierungsmaßnahmen“ hatten diese aber bis Mai 2017 durchschnittlich nur 21,6 Prozent der vorgesehenen Ausbildung durchlaufen (www.zeit.de/politik/deutschland/2017-06/bamf-ausbildungsquote-qualifikation-entscheider). Die Fragestellerinnen und Fragesteller befürchten vor diesem Hintergrund, dass die Entscheiderinnen und Entscheider möglicherweise nicht ausreichend geschult sind, um die Ergebnisse der softwarebasierten Analysen einzuordnen.

Zudem bestehen aus Sicht der Fragestellerinnen und Fragesteller Zweifel an der korrekten Anwendung der Spracherkennungssoftware durch die Mitarbeiterinnen und Mitarbeiter des BAMF. Laut einer internen Dienstanweisung des BAMF, die den Fragestellerinnen und Fragestellern vorliegt, soll der Antragsteller bzw. die Antragstellerin mindestens zwei Minuten möglichst unterbrechungsfrei sprechen. Ein Ergebnisbericht, der den Fragestellerinnen und Fragestellern vorliegt, dokumentiert jedoch lediglich eine Aufnahmedauer von 57,14 Sekunden und eine Nettosprachdauer von 25,7 Sekunden, womit die vorgesehene Mindestsprachdauer von zwei Minuten deutlich unterschritten wird.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Vorabfassung - wird durch die lektorierte Version ersetzt.

Wir fragen die Bundesregierung:

1. In wie vielen Fällen wurden in den Jahren 2016 und 2017 jeweils Sprach- und Textanalysegutachten im Auftrag des BAMF durchgeführt, um Angaben von Asylantragstellerinnen und -antragstellern zur Herkunft zu überprüfen?

   a) In wie vielen Fällen bestätigten die Gutachten die Angaben der Asylantragstellerinnen und Antragsteller zu ihrer Herkunft, in wie vielen Fällen brachten sie ein abweichendes Ergebnis?

   b) Wie wurde in den Fällen, in welchen die Angaben der Asylsuchenden von dem Ergebnis des Gutachtens abwichen, von den Entscheiderinnen und Entscheidern entschieden?

   c) Welche statistischen Angaben liegen zu den Ergebnissen dieser Gutachten darüber hinaus vor?

2. In wie vielen Fällen wurde die Dialekterkennungssoftware bereits eingesetzt (bitte nach BAMF-Standorten aufschlüsseln)?

3. In wie vielen Fällen wurden jeweils in den Jahren 2017 und 2018 bereits für die Prüfung der Angaben zur Staatsangehörigkeit in Auftrag gegebene, aber noch nicht durchgeführte Sprach- und Textanalysegutachten storniert, um statt dessen eine sprachbiometrische Analyse mittels Spracherkennungssoftware durchzuführen?

4. Welche Schulungsunterlagen und Interpretationshilfen werden den Mitarbeiterinnen und Mitarbeitern des BAMF zur Verfügung gestellt, um die Ergebnisberichte der Spracherkennungssoftware angemessen zu interpretieren?

5. Wie viele Mitarbeiterinnen und Mitarbeiter in den Außenstellen des BAMF wurden bereits als Multiplikatorinnen und Multiplikatoren ausgebildet, um den Umgang mit der Spracherkennungssoftware am Arbeitsplatz zu unterstützen (vgl. Bundestagsdrucksache 19/190, S. 7)?

6. Wer führt die zusätzlichen eintägigen Schulungsseminare zu den Assistenzsystemen durch, die laut Antwort der Bundesregierung auf die Kleine Anfrage der Fraktion DIE LINKE. bereit gestellt werden, um über die regulären Schulungen hinausgehende Schulungsbedarfe zu decken (Bundestagsdrucksache 19/190, S. 7)?

   a) Wie häufig werden diese zusätzlichen Seminare angeboten?

   b) Wie viele Entscheiderinnen und Entscheider bzw. Anhörerinnen und Anhörer haben an diesen Schulungen bereits teilgenommen?

   c) Wie viel Raum nimmt in den Schulungsseminaren der Umgang mit der Spracherkennungssoftware ein, wie viel Raum nimmt die Anwendung anderer Assistenzsysteme ein und welche sind dies?

   d) Findet eine Erfolgskontrolle statt und falls ja, wie?

7. Wie lässt sich erklären, dass die Spracherkennungssoftware des BAMF laut der Antwort der Bundesregierung auf die Kleine Anfrage der Fraktion DIE LINKE. vom 8. Dezember 2017 momentan nur für Arabisch in den Dialekten Ägyptisch, Irakisch, Levantinisch und Golf-Arabisch zur Verfügung steht (vgl. Bundestagsdrucksache 19/190, S. 3), in einer internen Dienstanweisung des BAMF, die den Fragestellerinnen und Fragestellern vorliegt, aber neben den bereits genannten Sprachen auch Bulgarisch, Deutsch, Englisch, Amerikanisch, Hebräisch, Hindi, Italienisch, Japanisch, Kanadisch Französisch, Koreanisch, Kroatisch, Latein-Amerikanisches Spanisch, Mandarin Chinesisch, Niederländisch, Paschtu, Farsi, Dari, Portugiesisch und Rumänisch angeführt werden?

8. Welche Erkenntnisse hat die Bundesregierung darüber, inwiefern die Spracherkennungssoftware von den Mitarbeiterinnen und Mitarbeitern des BAMF gemäß der internen Dienstanweisung angewendet wird?

   Wie stellt sie sicher, dass die Mitarbeiterinnen und Mitarbeiter des BAMF die Spracherkennungssoftware vorschriftsgemäß anwenden und Anwendungsfehler (siehe Vorbemerkung) gegebenenfalls korrigiert werden?

9. Inwiefern wird den Antragstellerinnen und Antragstellern die Möglichkeit gegeben, zu Widersprüchen zwischen ihren Angaben und dem Ergebnis des sprachbiometrischen Verfahrens Stellung zu nehmen?

10. In wie weit sind nach Kenntnis der Bundesregierung Ablehnungsbescheide zu Asylanträgen, die weitgehend auf den Ergebnissen der biometrischen Sprachanalyse fußen, oder solche Gutachten selbst Gegenstand von verwaltungsgerichtlichen Auseinandersetzungen?

11. Ist der Bundesregierung bekannt, ob im Zuge dessen von zuständigen Verwaltungsgerichten im Rahmen der Sachverhaltsaufklärung die Offenlegung der verwendeten algorithmischen bzw. statistischen Verfahren der Analysesoftware verlangt wurde?

12. Inwiefern werden Daten über in Nachrichten verwendete Sprachen, die bei der Auslese von Smartphones von Geflüchteten gewonnen werden, bei der Überprüfung von Angaben zum Herkunftsland genutzt?

    a) Falls ja, mithilfe welcher Verfahren werden die Daten ausgewertet?

       Inwiefern werden dabei Chat-Dialekte im Arabischen berücksichtigt?

    b) Falls nein, welche Erkenntnisse liegen der Bundesregierung zu Anbietern solcher Analyse-Verfahren oder -tools vor, und welche Maßnahmen plant sie zur Erforschung, Marktsichtung, Prüfung und ggf. Anschaffung solcher Verfahren oder Produkte?

13. Welche Kosten sind mit aktuellem Stand

    a) zur Anschaffung der hier in Rede stehenden biometrischen Sprachanalysesoftware,

    b) zur Anschaffung der Hard- und Software zum Auslesen von Mobiltelefonen, Smartphones und ähnlicher Endgeräte,

    c) zur Anschaffung von Hard- und Software zur Analyse der ausgelesenen Inhalte,

    d) zur Wartung und Pflege der in a)-c) genannten Hard- und Software,

    e) für Updates und Produkterweiterungen für die in a)-c) genannte Hard- und Software,

    f) für die oben genannten Qualifizierungsmaßnahmen für Mitarbeiterinnen und Mitarbeiter des BAMF

    entstanden und im laufenden Haushalt vorgesehen (bitte für die Jahre 2017 und 2018 getrennt angeben)?

Vorabfassung - wird durch die lektorierte Version ersetzt.

14. Welche Fehlerquellen bestehen aus Sicht der Bundesregierung bei der Anwendung der Spracherkennungssoftware?

    a) Inwiefern trägt die Spracherkennungssoftware der Tatsache Rechnung, dass sich die Sprachkompetenz über die Lebensspanne verändern kann (siehe Vorbemerkung)?

    b) Inwiefern können Hintergrundgeräusche auf die Analyseergebnisse Einfluss nehmen und wie wird diesem Problem ggf. begegnet?

    c) Wie wird damit umgegangen, wenn bei Asylantragstellerinnen und Asylantragstellern eine Sprachbehinderung wie etwa Lispeln oder Stottern vorliegt?

15. Welche Schlussfolgerungen zieht die Bundesregierung aus den bisherigen Erfahrungen mit dem Einsatz der Spracherkennungssoftware?

Berlin, den 22. März 2018

Dr. Sahra Wagenknecht, Dr. Dietmar Bartsch und Fraktion

Vorabfassung - wird durch die lektorierte Version ersetzt.

Satz: Satzweiss.com Print, Web, Software GmbH, Mainzer Straße 116, 66121 Saarbrücken, www.satzweiss.com
Druck: Printsystem GmbH, Schafwäsche 1-3, 71296 Heimsheim, www.printsystem.de
Vertrieb: Bundesanzeiger Verlag GmbH, Postfach 10 05 34, 50445 Köln, Telefon (02 21) 97 66 83 40, Fax (02 21) 97 66 83 44, www.betrifft-gesetze.de
ISSN 0722-8333